In the Claims

1 (currently amended). A method of operating a computer to identify all chemical structures defined by a Markush type formula (200, 220, 260), which is stored in a database matching a given query structure (200), without the necessity of generating said chemical structures, comprising the steps of:

(i) Processing said Markush type formula(e) and said query(ies) into a computer readable form (210),

(ii) Searching for partially relaxed subgraph isomorphism(s) for each said query (230, 240, 250), and

(iii) Retrieving data (270).

2 (original). The method of claim 1, wherein said database is made of at least one combinatorial library stored as a Markush type formula (200).

3 (original). The method of claim 2, wherein said libraries are each made of one scaffold and at least one R-group as constituents.

4 (original). The method of claim 1, wherein said given query structure is either an exact chemical structure or a chemical substructure.

5 (original). The method of claim 1, wherein said given query structure is said to match said chemical structure if said given query structure is exactly said chemical structure.

6 (original). The method of claim 1, wherein said given query structure is said to match said chemical structure if said given query structure is either said chemical structure or either a substructure of said chemical structure.

7 (original). The method of claim 1, wherein said identification can be performed with said query structure as sole input (200), without the requirement of additional information to perform said identification.

8 (original). The method of claim 1, wherein said generation of chemical structures is neither required before nor during the search.

9 (original). The method of claim 1, wherein said processing of said step (i) can either be performed before or either during said identification.

10 (original). The method of claim 1, wherein said Markush type formula(e) can either be pre-processed (210) or processed during said identification.

11 (original). The method of claim 1, wherein said query(ies) is(are) stored or not in a database.

12 (currently amended). The method of claim 3, wherein said processing of said step (i) comprises the steps of:

(a) Building of graphs and binary description of said scaffolds and R-groups, and

(b) Building of graph and binary description of said query(ies).

13 (currently amended). The method of claim 12, wherein said binary description of said step (a) contains or consists of the following information:

1. For each scaffold:
   (a) Number of atoms present in said scaffold,
   (b) Graph of said scaffold,
   (c) Number of R-groups,
   (d) Label of said R-groups,
   (e) Position of said R-groups in said graph,

(f) Number of neighbours for each R-group and position of said neighbours in said graph, and

2. For each R-group:

(a) R-group identification (ID),

(b) Number of atoms present in said R-group,

(c) Graph of said R-group,

(d) Number of attachment points in said R-group,

(e) Attachment points identification (atoms indexing),

(f) Atoms involved in said attachment points.

14 (currently amended). The method of claim 3, wherein said partially relaxed subgraph isomorphism searching of said step (ii) of said claim 1 (240) is performed on all said libraries and comprises the steps of:

(a) Scaffold reading (300),

(b) Partially relaxed subgraph isomorphism searching of said query against said scaffold (310), and

(c) Processing of all isomorphisms (320 to 390),

for each library of said database (220, 260).

15 (currently amended). The method of claim 14, wherein said processing of said step (c) comprises the step of:

(1) Counting the number of atoms of said query associated with each constituent of said library (330),

(2) Identifying which atoms of said query are associated with said constituent(s) (330),

(3) Identifying on which constituent(s) said query is located (330), and

(4) Processing of said isomorphism taking into account said query location of said step (3) (340 to 380),

for each isomorphism detected.

16 (original). The method of claim 15, wherein said step (3) defines the global localisation of said query on said library constituent(s) as being either only the scaffold (340), or either only one single R-group (350) or either the scaffold and at least one R-group (350).

17 (currently amended). The method of claim 15, wherein said processing of said step (4) comprises the steps of:

(i) Processing of said isomorphism if said query is only located on the scaffold of said library (370),

(ii) Processing of said isomorphism if said query is only located on a single R-group of said library (380),and

(iii) Processing of said isomorphism if said query is located on the scaffold and at least one R-group of said library (360 = all other cases).

18 (original). The method of claim 17, wherein said processing of said step (i) (370) comprises the step of storing said chemical structures of claim 1 matching the query as a sub-library identical to said library (400).

19 (currently amended). The method of claim 17, wherein said processing of said step (ii) (380) comprises the steps of:

(a) Identifying members of said single R-group containing said query (500, 510, 530, 700 to 730), and

(b) Flagging said members (520).

20 (original). The method of claim 19, wherein said chemical structures of claim 1 matching the query are stored as a sub-library corresponding to a Markush type formula made of said scaffold of claim 14, all members of R-groups not associated to said query and said flagged members of said single R-group identified by said query in said step (a) of claim 19 (550), if said single R-group has at least one member flagged (540).

21 (currently amended). The method of claim 17, wherein said processing of said step (iii) (360) comprises the steps of:

(a) Identifying if atoms of said query are associated with an R-group (610),

(b) Isomorphism searching (640, 700 to 730) of the sub-query (620) formed by said atoms, on each member (630, 660) of said associated R-group, if at least one atom is associated to said R-group (610), and

(c) Flagging each member of said associated R-group for which at least one isomorphism is detected (650),

for each R-group of said library (600, 670).

22 (original). The method of claim 21, wherein all members of an R-group of said library are flagged if said R-group is not involved in said isomorphism of step (b) of claim 14.

23 (original). The method of claim 21, wherein said chemical structures of claim 1 matching the query are stored as a sub-library corresponding to a Markush type formula made of said scaffold of claim 14, all members of R-groups not associated to said query and said flagged members of said associated R-groups (690), if all said associated R-groups have at least one member flagged (680).

24 (original). The method of claim 23, wherein said flagged members that match said sub-query are kept in a list for said isomorphism searching as IDs pointing to graphs.

25 (original). The method of claim 21, wherein the association of atoms in said query with atoms in said scaffold is saved, defining the partial localisation of said query on the sub-library.

26 (original). The method of claim 21, wherein a same list of members is used for different R-groups of said library sharing the same members.

27 (currently amended). The method of claim 21, wherein said sub-query isomorphism searching of said step (b) comprises the steps of:

(1) Building said sub-query to be searched in said associated R-group (620),

(2) Determining attachment point's constraints (620), and

(3) Isomorphism searching (640, 700 to 730) with said attachment points' constraints for each said associated R-group's member (630, 660).

28 (original). The method of claim 27, wherein graph connectivity of said sub-query is checked in step (1), meaning that atoms associated to a given R-group make a connected graph.

29 (original). The method of claim 27, wherein said isomorphism searching of said step (3) is partially relaxed or not (720 or 710).

30 (currently amended). The method of claim 27, wherein said determining of attachment points' constraints of said step (2) is defined as follows:

(i) For each neighbour C[i] of order i of said R-group in said scaffold, if said neighbour is associated to an atom of said query then D[i] represents said atom in said query, otherwise D[i]=∅,

(ii) For each said order i, if D[i] is defined then for each of the neighbour of D[i] in said query, if said neighbour is mapped to said R-group, A[i] represents said neighbour, otherwise A[i] is not defined (A[i]=∅), and

(iii) The array A represents the constraints of said attachment points.

31 (currently amended). The method of claim 27, wherein said isomorphism searching of said step (3) comprises the steps of:

(a) Reading said member (630), and

(b) Searching of all the isomorphisms of said sub-query (640, 700 to 730) on said member with said constraints on attachment points: said atom A[i] of said sub-query must be mapped to the attachment point of order i of said member, for each i where A[i] is defined.

32 (original). The method of claim 31, wherein the number of isomorphisms is counted in said step (b).

33 (currently amended). The method of ~~step~~ claim 31, wherein only the first isomorphism is searched in said step (b).

34 (original). The method of claim 31, wherein said method further comprises the step of saving all the isomorphism's descriptions, which defines, along with said partial localisation, the exact localisation of said query on said library.

35 (currently amended). The method of claim 31, wherein said searching of said isomorphisms (640, 700 to 730) of said step (b) comprises the additional steps of:

(i) Analysing each said member for the presence of a nested R-group (700), and

(ii) Proceeding recursively to claim 14 (720=240) with said query of said claim 14 corresponding now to said sub-query, said scaffold of said claim 14 corresponding now to said R-group and said R-groups are the said nested ones, until said nested R-groups are no more involved in an isomorphism, if said member contains a nested R-group (700).

36 (currently amended). The method of claim 3, wherein said data retrieval of said step (iii) retrieves at least one of the following information:

- For the entire said database:
  - Said database contains or does not contain said query or is there at least one said library that contains said query,
  - A list of all the combinatorial libraries containing said query,
  - A list of all the combinatorial libraries not containing said query,
  - A list and number of said scaffolds containing entirely said query,
  - A list and number of said scaffolds not containing entirely said query,

- A list and number of said R-groups containing entirely said query whether nested R-groups are allowed or not,
- A list and number of said R-groups not containing entirely said query whether nested R-groups are allowed or not,
- The total number of isomorphisms retrieved in step (b) of claim 14 (310) for all the libraries, whether said associated R-groups of claim 15 have at least one member flagged during said processing of said step (4) or not (540, 680),
- The global or partial localisation for all the isomorphisms,
- The first isomorphism found with or without its global or partial localisations,

- For each said library:
  - Said library contains or does not contain said query,
  - A list and number of all the enumerated (specific) structures or non-enumerated structures of said library matching said query,
  - The number of unique structures of said library matching said query, whatever the number of partial localisations of said query on said library,
  - The number of times said query is located on said scaffold only, or on said R-groups only, or spans across said scaffold and said R-group(s). This corresponds to the number of global localisations,
  - The total number of isomorphisms retrieved in step (b) of claim 14, whether said associated R-groups of claim 15 have at least one member flagged during said processing of step (4) or not. This corresponds to the total of the number of said partial localisations of said query on said library,
  - A list of all said partial localisations of said query on said library, each one corresponding to an isomorphism and defining a sub-library,
- For each said R-group:
  - Said R-group contains or does not contain said query or said sub-query,

- A list and number of all the specific members or non-enumerated members of said R-group containing said query or said sub-query, whether nested R-groups are allowed or not,
- A list and number of all the specific members or non-enumerated members of said R-group not containing said query or said sub-query, whether nested R-groups are allowed or not,
- The number of times said query or sub-query is found in said R-group's members whether exact localisation or nested R-groups are taken into account or not. This corresponds to the total number of isomorphisms for all said R-group's members,

- For each said member of said R-group:
  - Said member contains or does not contain said query or said sub-query,
  - The number of times said query or sub-query is found on said member whether nested R-groups are taken into account or not. This corresponds to the number of isomorphisms of said sub-query on said member,
  - A list and number of all the specific structures or non-enumerated structures described by said member containing said query or said sub-query if said member contain nested R-group(s),
  - The exact localisation of said query or sub-query on said member,
- For each single isomorphism of said query or sub-query:
  - The library corresponding to said isomorphism,
  - A list and number of R-groups associated to at least one atom of said query in said isomorphism,
  - A list and number of R-groups not associated to any of the atoms of said query in said isomorphism,
  - A list and number of members containing said query or said sub-query for each said R-group,
  - A list and number of members not containing said query or said sub-query for each said R-group,

- The global localisation of said query on said library, i.e. said query is either only on the scaffold, or either only on one R-group or either on the scaffold and at least one R-group,
- The partial localisation of said query on said library, i.e. the atoms in the scaffold and the R-group(s) to which atoms in the query are mapped,
- A list of all the specific structures or non-enumerated structures containing said query and mapping on said library following said partial localisation, and

• For all the isomorphisms of said query or sub-query:

- All the information gathered in the aforementioned points.

37 (currently amended). The method of ~~any of claims 1 or 36~~ claim 1, wherein said data retrieval of said step (iii) retrieves said structures in the form of either enumerated or either non-enumerated structures.

38 (currently amended). The method of ~~any of claims 1 or 36~~ claim 1, wherein said data retrieval of said step (iii) takes into accounts nested R-groups.

39 (currently amended). The method of ~~any of claims 1 or 36~~ claim 1, wherein said data retrieval of said step (iii) takes into account the exact localisation of said query for each said isomorphism.

40 (currently amended). The method of ~~any of the preceding claims, wherein~~ claim 1, further comprising the application of screening technique(s) option(s) ~~is applied, thereby reducing searching~~ to reduce search time.

41 (currently amended). The method of claim 40, wherein said screening technique option relies on substructural features ~~such as keys~~.

42 (currently amended). The method of ~~any of the preceding claims~~ claim 1, wherein it can be integrated in a pipeline.

43-48 (canceled).

49 (currently amended). ~~Drug~~ A drugcompound obtained by synthesising a molecule determined by performing the method according to ~~any of the preceding claims~~ claim 1.

50 (new). A computer program for the automatic identification of all the chemical structures defined by a Markush type formula(e), which is stored in a database matching a given query structure, without the necessity of generating said chemical structures, comprising computer code means adapted to perform the method of claim 1 when said program is run on a computer; or

computer readable medium having a program recorded thereon, said program comprising a computer program for the automatic identification of all the chemical structures defined by a Markush type formula(e), which is stored in a database matching a given query structure, without the necessity of generating said chemical structures, comprising computer code means adapted to perform the method of claim 1 when said program is run on a computer;

a computer loadable product directly that is loadable into the internal memory of a digital computer, said product comprising software code portions the automatic identification of all the chemical structures defined by a Markush type formula(e), which is stored in a database matching a given query structure, without the necessity of generating said chemical structures, adapted to perform the method of claim 1 when said product is run on a computer; or

an apparatus for performing the method of claim 1, said apparatus comprising data input means for inserting said at least one given query structure.